CLAIMS

1. DNA-binding molecule, capable of sequence specific binding to the minor groove of double-stranded DNA, characterised in that it comprises at least two sequence specific DNA-binding elements, covalently linked to each other in tandem orientation by an amphipathic, flexible linker molecule, at least one of said DNA binding elements being non-proteinaceous.

2. DNA-binding molecule according to claim 1 wherein at least one of the DNA-binding elements comprises an oligomer comprising one or more organic heterocyclic amino-acid residues.

3. DNA-binding molecule according to claim 2 wherein each organic heterocyclic residue has at least one annular nitrogen, sulphur or oxygen.

4. DNA-binding molecule according to claim 2 or 3 wherein said heterocyclic residue is chosen from pyrrole, imidazole, triazole, pyrazole, furan, thiazole, thiophene, oxazole, pyridine, or derivatives of any of these compounds wherein one or more of the heteroatoms are substituted by a substituent which is DNA-binding or non-DNA-binding.

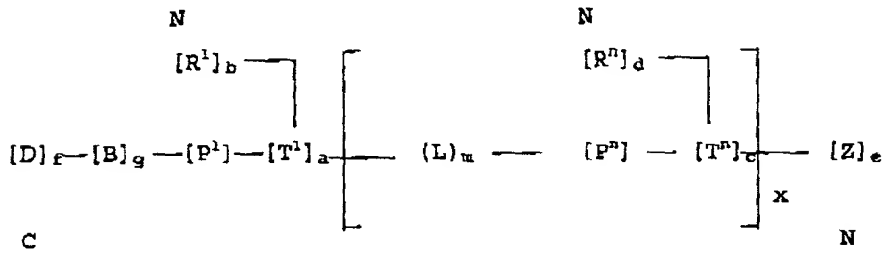
5. DNA-binding molecule according to claim 4 wherein at least one oligomer includes heterocyclic residues chosen from N-methylpyrrole (Py) and /or 3-hydroxy N-methylpyrrole (HP) and / or N-methylimidazole (Im).

6. DNA-binding molecule according to any one of claims 2 to 5 wherein the DNA-binding element further comprises at least

one aliphatic amino acid residue such as a β-alanine (β) residue, or a 5-aminovaleric acid residue.

7. DNA-binding molecule according to any one of claims 1 to 4, having the general formula (I) :

(I)



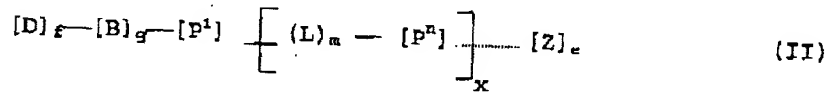
wherein

- each of P^1 to P^n represents a DNA-binding element, said element comprising multiple organic heterocyclic or aliphatic residues or fluorescent derivatives thereof ;
- each of R^1 to R^n represents a DNA-binding element, said element comprising multiple organic heterocyclic or aliphatic residues or fluorescent derivatives thereof ;
- x represents an integer from 1 to 20, with the proviso that when x is greater than 1, the multiple copies of $[R^n]$, $[L^n]$, $[P^n]$ and $[T^n]$ may be the same or different ;
- n represents an integer having a value equal to (x+1) ;
- [T] represents a multifunctional linking molecule providing a covalent link between DNA-binding elements [R] and [P], with the proviso that if "e" represents 0, $[T^{x+1}]$ can be bifunctional;
- each of a and c independently represent 0 or 1 ;

- each of b and d independently represent 0 or 1, with the proviso that when a represents 0, b also represents 0, and when c represents 0, d also represents 0 ;
- [D] represents an end group or an effector moiety,
- $[L]_m$ represents an amphipathic, flexible linker molecule, linking the DNA-binding elements in a tandem orientation with respect to each other ;
- m represents an integer from 1 to 10 ;
- [B] represents a spacer unit such as β-alanine ;
- [Z] represents an end group or an effector moiety ;
- each of f, g and e independently represent 0 or 1,
- each solid line represents a covalent bond ;
- N and C indicate the N- and C-terminal extremities of the molecule, respectively.

8. DNA-binding molecule according to claim 7 wherein the DNA-binding elements P and R comprise heterocyclic residues chosen from pyrrole, imidazole, triazole, pyrazole, furan, thiazole, thiophene, oxazole, pyridine, or derivatives of any of these compounds wherein one or more of the heteroatoms is substituted.

9. DNA-binding molecule according to claim 8 having the general formula (II) :



(II)

wherein $[P^1]$, $[P^n]$, (L), [D], [Z], x, m, f, g and e have the previously defined meanings
and a dotted line represents a covalent bond which can be present or absent.

10. DNA-binding molecule according to claim 9 wherein each of the the DNA-binding elements $[P^1]$ to $[P^n]$ independently have the general formula (III)

$$---[-U^1 - [U]_s]---\qquad (III)$$

wherein :
each U is independently a monomeric unit chosen from a heterocyclic amino acid residue, or an aliphatic amino acid residue or a fluorescent derivative thereof, and
s is an integer from 1 to 15, preferably from 2 to 8,
and a dotted line represents a covalent bond which can be present or absent.

11. DNA-binding molecule according to claim 10 wherein at least one U is chosen from N-methylpyrrole (Py) and /or 3-hydroxy N-methylpyrrole (HP) and / or N-methylimidazole (Im).

12. DNA-binding molecule according to claim 10 wherein at least one U is a β-alanine (β) residue, or a 5-aminovaleric acid residue.

13. DNA-binding molecule according to claim 11 or 12 wherein S is an integer from 2 to 5

14. DNA-binding molecule according to claim 13 wherein at least one of $[P^1]$ to $[P^n]$ comprises between 3 to 5 heterocyclic amino acid residues.

15. DNA-binding molecule according to claim 13 wherein at least one of $[P^1]$ to $[P^n]$ comprises more than two contiguous

heterocyclic amino acid residues, for example three, four or five contiguous heterocyclic amino acid residues.

16. DNA-binding molecule according to claim 15 wherein stretches of three to five contiguous heterocyclic amino acid residues are separated from each other by a β-alanine residue

17. DNA-binding molecule according to claim 10 wherein at least one of $[P^1]$ to $[P^n]$ has the formula (IV)

N C

$$\cdots\left[[U_1]-[U_2]-[U_3]-[U_4]-[U_5]-[U_6]-[U_7]-[U_8]-\right]\cdots \qquad \text{(IV)}$$

wherein U is as previously defined,
$[U_4]$ is β-alanine,
$[U_1]$ to $[U_3]$, and $[U_5]$ to $[U_7]$ are chosen from N-methylpyrrole (Py) and / or N-methylimidazole (Im),
$[U_8]$ may be present or absent, and if present is preferentially β-alanine,
and a dotted line represents a covalent bond which can be present or absent.

18. DNA-binding molecule according to claim 17 wherein $[U_1]$ to $[U_3]$, and $[U_5]$ to $[U_7]$ are each N-methylpyrrole (Py).

19. DNA-binding molecule according to claim 10 wherein at least one of $[P^1]$ to $[P^n]$ has the fomula (V) :

N C

$$\cdots\left[[U_1]-[U_2]-[U_3]-[U_4]-[U_5]-[U_6]-[U_7]-[U_8]-[U_9]\right]\cdots \qquad \text{(V)}$$

wherein :

- U is as previously defined,
- $[U_1]$ to $[U_8]$ are chosen from N-methylpyrrole (Py), N-methylimidazole (Im) and a β alanine residue,
- with the proviso that the [U] immediately adjacent to each Im on the N-terminal side is a β alanine residue, $[U_9]$ may be present or absent, and if present is preferentially β-alanine,

and a dotted line represents a covalent bond which can be present or absent.

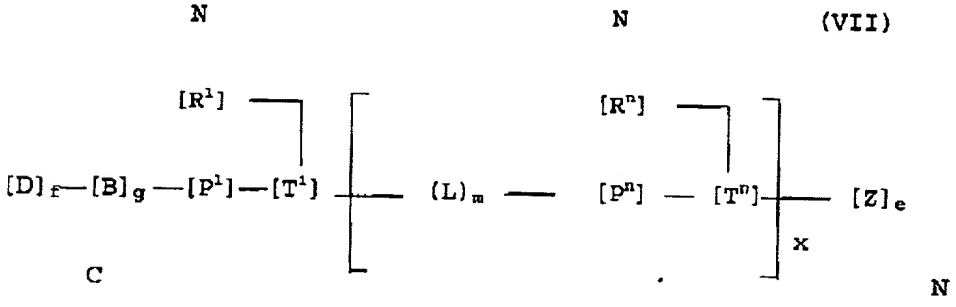
20. DNA-binding molecule according to claim 19 wherein at least one of $[P^1]$ to $[P^n]$ has the formula (VI) :

N C

.....[Im-β-Im-Py-β-Im-β-Im-β]..... (VI)

21. DNA-binding molecule according to any one of claims 9 to 20 wherein x represents a value from 2 to 10, for example 2, 3, 4, 5, 6, 7, 8, 9, or 10.

22. DNA-binding molecule according to claim 7 having the general formula (VII) :



wherein $[R^1]$, $[P^1]$, $[R^n]$, $[P^n]$, $[T^1]$, $[T^n]$, (L), [D], [B], [Z], m, n, g, f and e have the previously defined meanings.

23. DNA-binding molecule according to claim 22 wherein each of the DNA-binding elements $[P^1]$ to $[P^n]$ and $[R^1]$ to $[R^n]$ independently have the general formula (VIII)

$$-\cdots[U^1 - [U]_s-]- \qquad \text{(VIII)}$$

wherein :

each U is independently a monomeric unit chosen from a heterocyclic amino acid residue, or an aliphatic amino acid residue or a fluorescent derivative of the foregoing, and s is an integer from 0 to 15, preferably from 1 to 6 and a dotted line represents a covalent bond which can be present or absent.

24. DNA-binding molecule according to claim 23 wherein at least one heterocyclic amino acid residue comprises an annular nitrogen.

25. DNA-binding molecule according to claim 24 wherein at least one of $[P^1]$ to $[P^n]$ or $[R^1]$ to $[R^n]$ contain a residue of N-methylpyrrole (Py) and /or 3-hydroxy N-methylpyrrole (HP) and / or N-methylimidazole (Im).

26. DNA-binding molecule according to claim 25 wherein at least one of $[P^1]$ to $[P^n]$ or $[R^1]$ to $[R^n]$ contain an aliphatic amino-acid residue such as a β-alanine (β) residue.

27. DNA-binding molecule according to claim 23 or 24 wherein S is an integer from 2 to 6.

28. DNA-binding molecule according to claim 23 or 24 wherein at least one of $[P^1]$ to $[P^n]$ or $[R^1]$ to $[R^n]$ comprises at least four heterocyclic amino acid residues.

29. DNA-binding molecule according to claim 23 wherein at least one of $[P^1]$ to $[P^n]$ or $[R^1]$ to $[R^n]$ comprises more than two contiguous heterocyclic amino acid residues, for example three, four or five contiguous heterocyclic amino acid residues.

30 DNA-binding molecule according to claim 29 wherein stretches of three to five contiguous heterocyclic amino acid residues are separated from each other by a β-alanine residue.

31. DNA-binding molecule according to claim 23 wherein at least one $[P^n]$ element has the formula (IX) :

N C

$$[[U_1]-\ldots-[U_{s-1}]-[U_s]] \qquad (IX)$$

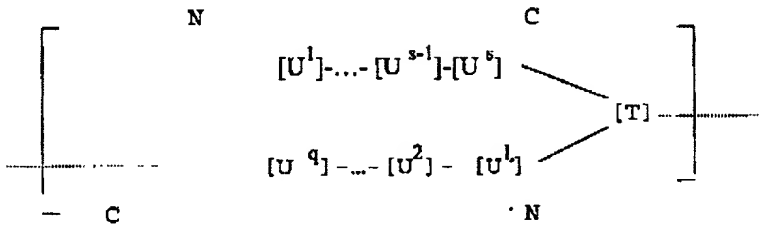
and at least one $[R^n]$ element has the formula (X) :

N C

$$[[U_q]-\ldots[U_2]-[U_1]] \qquad (X)$$

wherein each U represents independently N-methylpyrrole (Py), or 3-hydroxy N-methylpyrrole (HP), or N-methylimidazole (Im) or N-methyl pyrazole (Pz), or 3-pyrazolecarboxylic acid

(3-Pz), or β-alanine (β), q and s are independently integers from 1 to 10,
and a dotted line represents a covalent bond which can be present or absent,
wherein the U residues of [P^n] form anti-parallel pairs with the U residues of [R^n] :



said pairs being chosen from Py/Im, Im/Py, Py/Py, Hp/Py, Py/Hp, β/Py, Py/β, β/Im, Im/β, Im/Im, Pz/Py, 3-Pz/Pz, and β/β.

33. DNA-binding molecule according to claim 9 wherein at least one DNA-binding element contains [T], [R] and [P] moieties, and at least one DNA binding element is free of [T] and [R] moieties.

34. DNA-binding molecule according to claim 9 wherein the multiple [R] and [P] elements are different in length and / or composition.

35. DNA-binding molecule according to any one of the preceding claims 7 to 34 having the capacity to bind in a multidentate mode to a given strand of DNA.

36. DNA-binding molecule according to any one of the preceding claims 7 to 35 wherein m represents a value greater than or

equal to one, and the amphipathic linker $(L)_m$ comprises an assembly of linker sub-units (L).

37. DNA-binding molecule according according to claim 36 wherein the assembled linker $(L)_m$ is heterobifunctional.

38. DNA-binding molecule according to claim 36 wherein each linker sub-unit (L) is heterobifunctional.

39. DNA-binding molecule according to claim 36, wherein at least one (L) sub-unit is amphipathic.

40. DNA-binding molecule according to claim 37 wherein the total length of the linker $(L)_m$ is between 5 to 250 Angstroms, for example 5 to 50 Angstroms.

41. DNA-binding molecule according to claim 37 wherein the functional groups are chosen from amino, carboxyl, thiol, haloacetyl, aldehyde, amino-oxy, maleimide groups, a symetrical anhydride and halogen atoms.

42. DNA-binding molecule according to claim 36 wherein at least one amphipathic linker (L) sub-unit comprises one or more ether groups and/or ester groups for example molecules derived from ethylene oxide or propylene oxide.

43. DNA-binding molecule according to claim 36 wherein at least one amphipathic linker (L) sub-unit comprises one or more units of 8-amino-3,6-dioxaoctanoic acid (Ao).

44. DNA-binding molecule according to any one of claims 1 to 43 having the capacity to bind in a sequence specific manner to a DNA recognition sequence of at least 6, preferably at

least 10 and most preferably at least 14 base pairs in length.

45. DNA-binding molecule according to any one of claims 1 to 44 having a molecular weight no greater than approximately 8 kDa.

46. DNA-binding molecule according to any one of claims 1 to 45 wherein the said molecule binds to the DNA minor groove.

47. DNA-binding molecule according to any one of claims 1 to 46 which is cell-permeable.

48. DNA-binding molecule according to any one of claims 1 to 47 having an apparent binding affinity of at least 5×10^7 M^{-1}.

49. DNA-binding molecule according to any one of claims 1 to 48 having an apparent binding affinity of at least 1×10^9 M^{-1}.

50. DNA-binding molecule according to any one of claims 1 to 49 having an apparent binding affinity of at least 5×10^{10} M^{-1}.

51. Process for binding double-stranded DNA in a sequence-specific manner, comprising contacting a DNA-target sequence within said DNA with a DNA-binding molecule according to any one of claims 1 to 50, in conditions allowing said binding to occur.

52. Process according to claim 51 which is carried out *in vivo*, *in vitro* or *ex vivo*.

53. Process according to claim 52 which is carried out in a cell.

54. Process according to claim 53, wherein said cell is eukaryotic.

55. Process according to claim 53, wherein said cell is prokaryotic.

56. Process according to claim 54, wherein said cell is a vertebrate cell, an invertebrate cell, a plant cell

57. Process according to claim54, wherein said cell is a mammalian cell, an insect cell, or a yeast cell.

58. Process according to any one of claims 51 to 57 wherein the double stranded DNA is endogenous to said cell.

59. Process according to any one of claims 51 to 57 wherein the double stranded DNA is heterologous to said cell.

60. Process according to claim 53 wherein the double stranded DNA target sequence comprises a chromatin element.

61. Process according to claim 60 wherein the target sequence comprises a SAR-like sequence.

62. Process according to claim 60 wherein the target sequence comprises a GAGAA repeat sequence.

63. Process according to any one of claims 50 to 62 wherein the target sequence has at least 8 and preferably at least 15 bases.

64. Process according to claim 60 wherein the target sequence is a cis- or trans-acting element mediating chromosome function.

65. Process according to claim 64 wherein the binding of the target element with the sequence-specific binding molecule gives rise to cis- and / or trans-regulation of chromosome function.

66. Process according to claim 53 wherein the double stranded DNA target sequence comprises a site mediating the activity of one or more regulatory factors.

67. Process according to claim 66 wherein the regulatory factors is a transcription regulatory factor, a DNA replication factor, a factor for enzymatic activity, a factor involved in chromosome stability.

68. Process according to any one of claims 51 to 67, wherein the DNA-binding molecule is linked to an effector moiety.

69. Process for modulating chromosome function in a eukaryotic cell,
comprising the step of contacting a genomic DNA element, comprising a binding site mediating chromosome function,
with a molecule according to any one of claims 1 to 50 and having the capacity to bind in a sequence-specific manner to said element,
said step of contacting being carried out in conditions permitting binding of said compound to said element,
wherein the binding modulates chromosome function.

70. Process for modulating the function of a DNA element in a eukaryotic cell,
Comprising the step of contacting a genomic DNA element, so-called « chromatin responsive element » (CRE),
with a molecule according to any one of claims 1 to 50 and having the capacity to bind in a sequence-specific manner to said CRE,
said step of contacting being carried out in conditions permitting chromatin remodeling of the CRE by said compound,
wherein said chromatin remodeling of the CRE alters the activity of one or more other DNA elements, so called « modulated DNA elements » in the genome.

71. Cell containing a compound according to any one of claims 1 to 50.

72. Cell according to claim 71, wherein said compound binds the DNA-minor groove.

73. Cell according to claim 71 which is a eukaryotic cell.

74. Non-human organism comprising a cell according to claim 71.

75. Organism according to claim 74 which is a non-human animal.

76. Organism according to claim 75 which is a transgenic, non-human animal.

77. Organism according to claim 74 which is a plant.

78. Organism according to claim 77 which is a transgenic plant.

79. Pharmaceutical composition comprising a compound according to any one of claims 1 to 50 in association with a physiologically acceptable excipient.

80. Compound according to any one of claims 1 to 50, for use in therapy.

81. Compound according to any one of claims 1 to 50 which is fluorescent or fluorescently labelled.

82. Compound according to claim 81 wherein the fluorescent label is a fluorescent dye such as fluorescein, dansyl, Texas red, isosulfan blue, ethyl red, malachite green, rhodamine and cyanine dyes.

83. Use of a compound according to claim 81 for probing the epigenetic state and location of DNA in chromosomes and nuclei.

84. Use according to claim 83 for diagnosis of pathological conditions arising from epigenetic status.

85. Use of a compound according to claim 81 for chromosome visualisation and marking in diagnosis, forensic studies, affiliation studies, or animal husbandry.